07-03-00

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1 53(b))

Attorney Docket No.	1999-0260	Total Pages	54
First Named Inventor or Application Identifier			
Edmund Gale Acker et al.			
Express Mail Label No.		EL487593098US	

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO: Assistant Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

1. ☒ Fee Transmittal Form
   (submit an original, and a duplicate for fee processing)
2. ☒ Specification   [Total Pages 38]
   (preferred arrangement set forth below)
   - Descriptive title of invention
   - Cross References to Related Applications
   - Statement Regarding Fed sponsored R&D
   - Reference to Microfiche Appendix
   - Background of the Invention
   - Brief Summary of the Invention
   - Brief Description of the Drawings(if filed)
   - Detailed Description
   - Claim(s)
   - Abstract of the Disclosure
3. ☒ Drawing(s)(35 USC 113) [Total Sheets 5]
4. Oath or Declaration   [Total Pages 3]
   a. ☒ Newly executed (original or copy)
   b. ☐ Copy from a prior application(37 CFR 1.63(d))
      (for continuation/divisional with Box 17 completed)
      [Note Box 5 below]
      i. ☐ DELETION OF INVENTOR(S)
         Signed statement attached deleting inventor(s)
         named in the prior application, see 37 CFR
         163(d)(2) and 1.33(b)
5. ☐ Incorporation by reference(useable if Box 4b is checked)
   The entire disclosure of the prior application, from which a copy of the oath
   or declaration is supplied under Box 4b, is considered as being part of the
   disclosure of the accompanying application and is hereby incorporated by
   reference herein.

6. ☐ Microfiche Computer Program (Appendix)
7. ☐ Nucleotide and/or Amino Acid Sequence Submission
   (if applicable, all necessary)
   a. ☐ Computer Readable Copy
   b. ☐ Paper Copy(identical to computer copy)
   c. ☐ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. ☒ Assignment Papers(cover sheet & document(s))
9. ☐ 37 CFR 3.73(b)Statement ☐ Power of Attorney
10. ☐ English Translation Document (if applicable)
11. ☒ Information Disclosure ☒ Copies of IDS
    Statement(IDS)/PTO-1449   Citations
12. ☐ Preliminary Amendment
13. ☒ Return Receipt Postcard (MPEP 503)
    (Should be specifically itemized)
14. ☐ Small Entity ☐ Statement filed in prior application,
    Statement(s)   Status still proper and desired
15. ☐ Certified Copy of Priority Document(s)
    (if foreign priority is claimed)
16. ☐ Other :

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information:
☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior Application No:

18. CORRESPONDENCE ADDRESS

☐ Customer Number or Bar Code Label   (Insert Customer No. or Attach bar code label here)   or ☒ Correspondence address below

NAME	Samuel H. Dworetsky				
ADDRESS	AT&T CORP. P.O. Box 4110				
CITY	Middletown	STATE	New Jersey	ZIP CODE	07748-4110
COUNTRY	United States of America			FAX	732-368-6932

19. SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

NAME	Michele L. Conover		Reg. #	34962
TELEPHONE	908-221-5773			
SIGNATURE		DATE	June 30, 2000	

"Express Mail" Mailing Label Number   EL487593098US

Date of Deposit 6/30/2000

I hereby certify that this application is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington D.C. , 20231

Ann E. Taylor
(Printed Name of Person Mailing Paper)

(Signature of Person Mailing Paper)

PERSONALIZED TEXT-TO-SPEECH SERVICES

FIELD OF THE INVENTION

The present invention relates to text-to-speech conversion, and, more particularly,

5   is directed to services using a template for personalized text-to-speech conversion.

BACKGROUND OF THE INVENTION

Text-To-Speech (TTS) systems for converting text into synthesized speech are

entering the mainstream of advanced telecommunications applications. A typical TTS system

10   proceeds through several steps for converting text into synthesized speech. First, a TTS system

may include a text normalization procedure for processing input text into a standardized format.

The TTS system may perform linguistic processing, such as syntactic analysis, word

pronunciation, and prosodic prediction including phrasing and accentuation. Next, the system

performs a prosody generation procedure, which involves translation between the symbolic text

15   representation to numerical values of a fundamental frequency, duration, and amplitude.

Thereafter, speech is synthesized using a speech database or template comprising concatenation

of a small set of controlled units, such as diphones. Increasing the size and complexity of the

speech template may provide improved speech synthesis. Examples of TTS systems are

described in U.S. Patent No. 6,003,005, entitled "Text-To-Speech System And A Method And

20   Apparatus For Training The Same Based Upon Intonational Feature Annotations Of Input Text",

and U.S. Patent No. 5,774,854, entitled "Text To Speech System", which are hereby

incorporated by reference. Additional information about TTS systems may be found in "Talking

1

Machines: Theories, Models and Designs", ed G. Bailly and C. Benuit, North Holland (Elsevier), 1992.

SUMMARY

5   In accordance with an aspect of this invention, there are provided a method of and a system for providing services using a template for personalized text-to-speech conversion.

In general, in a first aspect, the invention features a method for converting text to speech, including receiving data representing a textual message that is directed from an author to a recipient, receiving information identifying an individual, retrieving a speech template 10 comprising information representing characteristics of the individual's voice, and converting the data representing the textual message to speech data. The speech data represents a spoken form of the textual message having the characteristics of the individual's voice.

In a second aspect, the invention features a text to speech conversion system, including a memory that stores executable program code, a processor that executes the program 15 code, and a storage device that stores a speech template comprising information representing characteristics of the individual's voice. The individual is identified by identification data. The program code is executable to convert text data to speech data. The text data represents a textual message directed from an author to a recipient, and the speech data represents a spoken form of the text data having the characteristics of the individual's voice.

20   In a third aspect, the invention features an article of manufacture including a computer readable medium having computer usable program code embodied therein. The computer usable program code contains executable instructions that when executed, cause a computer to perform the methods described herein.

2

515359_1

In a fourth aspect, the invention features a method for generating speech data for a voice response system, including receiving input from a recipient, generating a text message that provides a response to the input, selecting a speech template comprising information representing characteristics of a voice based at least in part on attributes of the recipient such as age or gender, and converting the text message to speech data. The speech data represents a spoken form of the textual message having the characteristics of the voice.

In a fifth aspect, the invention features a method for converting chat room text to speech, including storing a plurality of speech templates, each speech template comprising information representing characteristics of a chat room participant's voice, receiving the chat room text from an author who is a chat room participant, retrieving a speech template comprising information representing characteristics of the author's voice from the plurality of speech templates, and converting the chat room text to speech data. The speech data represents a spoken form of the textual message having the characteristics of the author's voice.

In a sixth aspect, the invention features a method for providing spoken electronic mail, including receiving an electronic text message addressed to a recipient from an author of the message, retrieving a speech template comprising information representing characteristics of the author's voice, converting the text message to speech data representing a spoken form of the textual message having the characteristics of the author's voice, and directing the speech data to the recipient.

In a seventh aspect, the invention features a method for providing speech output from a software application, including receiving text data from the software application, receiving information identifying an individual, retrieving a speech template comprising information representing characteristics of the individual's voice, converting the text data to

3

speech data representing a spoken form of the text data having the characteristics of the

individual's voice, and supplying the speech data to an output device for output to a user as audio

information. The software application may comprise an interactive learning program.

Preferred embodiments of the invention additionally feature the author interacting

5   with a first computer and the recipient interacting with a second computer which is coupled to

the first computer through a data network. The speech template may be provided at a central

location coupled to the first and second computers. Text data may be received at the central

location from either the first or second computer, and the speech data may be transmitted to the

first or second computer from the central location. Alternatively, the speech template may be

10   provided at the first computer, and either the speech data or the speech template may be

transmitted to second computer from the first computer. Alternatively, the speech template may

be provided at the second computer, and the data representing the textual message may be

received at the second computer.

In other embodiments, the first and second computers may communicate in an

15   instant messaging format, or they may be coupled to a server configured to operate chat room

software, with the text data comprising text input to the chat room. The server may store speech

templates for users of the chat room. The first and second computers may be coupled to a server

adapted to store and provide access to a shared space object that is associated with the textual

message. The data representing the textual message may also be an e-mail message.

20   In other embodiments, the recipient interacts with a telephone coupled to a

telephone network, and the author interacts with a computer coupled to the telephone network

through a data network. Input from the recipient may comprise telephone key depression or

speech. The speech data may be directed to the telephone network through the data network. A

notification may be transmitted to the author when the recipient is unable to connect with a telephone of the author, and the text data may be received in response to the notification message.

In other embodiments, the author may be defined as executable program code designed to generate text in response to input from the recipient. The individual may be selected based on attributes of the recipient, such as age or gender. The data representing the textual message may comprise a variable portion of a message having both a variable portion and a fixed portion, and it may further include the fixed portion. The fixed portion may be prerecorded speech of the individual.

It is not intended that the invention be summarized here in its entirety. Rather, further features, aspects and advantages of the invention are set forth in or will be apparent from the following description and drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a flow chart illustrating an embodiment for a personalized text-to-speech (pTTS) system;

Fig. 2 is an illustration of a pTTS system embodied in a stand alone personal computer;

Fig. 3 is an illustration of a pTTS system wherein a pTTS template associated with an author of a text message is stored on a centralized server;

Fig. 4 is an illustration of a pTTS system wherein a pTTS template associated with an author of a text message is stored on the author's computer;

515359_1

Fig. 5 is an illustration of a pTTS system wherein a pTTS template associated with an author of a text message is stored on a recipient's computer;

Fig. 6 is an illustration of a pTTS system wherein the server is coupled to a public switched telephone network; and

5   Fig. 7 is an illustration of a Chat implementation architecture.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

According to an embodiment of the present invention, a personalized text-to-speech (pTTS) system provides text-to-speech conversion for use with various services. These services, discussed in detail below, include, but are not limited to, speech announcements, film dubbing, Internet person-to-person spoken messaging, Internet chat room spoken text, spoken electronic mail, Internet shared spaces having objects intended for spoken presentation, and spoken notice of an incoming telephone call to a subscriber using the Internet.

Fig. 1 is a flowchart representing an embodiment for a pTTS system. In step 100, the pTTS system receives text data directed from an author of the text data to an intended recipient. The text data is provided in a data format representing a generic text message, such as a text file or a word processing file. In one embodiment, the recipient may be a specific person or group of people. For example, the text data may be an e-mail message sent by the author. Alternatively, the recipient may be unknown to the author. For example, the author may post the text data on a web site for access by unspecified users.

In step 102, the pTTS system identifies the author of the text data for enabling identification of the proper pTTS template. In one embodiment, the pTTS system identifies the author using the author's e-mail address. Alternatively, the pTTS system requests confirmation

6

of the author's identification by taking advantage of a user identification and/or password. In another alternative embodiment, the author's identification is transmitted with the text data in a predefined format. The identification step may additionally serve as an authentication or authorization step, to prevent unauthorized access to saved pTTS templates.

5   After the pTTS system identifies the author, the pTTS system retrieves a stored speech template associated with the author (step 104), referred to herein as the author's pTTS template. The author's pTTS template is a data file containing information representing voice characteristics of the author or voice characteristics selected by the author. Multiple pTTS templates are stored in the pTTS system for utilization by different users. In an alternative

10   embodiment, the pTTS system provides the author with the option to generate a new pTTS template, using methods known in the art. In another alternative embodiment, an author has more than one pTTS template, representing different types of speech or different voice characteristics. For example, an author provides pTTS templates having speech characteristics corresponding to different languages. An author having multiple pTTS templates selects the

15   appropriate pTTS template for the applicable text data. Alternatively, the author may have more than one user identification for accessing the pTTS system, each associated with a different pTTS template.

After retrieving the author's pTTS template, the pTTS system generates speech data (step 106) corresponding to the text data. The pTTS system takes advantage of the author's

20   pTTS template to generate the speech data in a format that may be audibly reproduced having voice characteristics represented by the selected template. For example, the speech data may be represented by data in the format of a standard ".wav" file. Thereafter, the speech data is output from the pTTS system (step 108), and transmitted to the appropriate destination.

7

515359_1

Referring to Fig. 2, stand alone personal computer 110 has memory 112 and storage 114, such as magnetic, optical, or magneto optical storage. Storage 114 includes at least one pTTS template 116. Personal computer 110 is programmed to select an appropriate pTTS template, which may be based on various factors, such as attributes of the author or recipient of

5   the message. Conversion routine 118 executing in memory 112 accepts text data and converts the text data to speech data with pTTS template 116, following the procedure outlined in Fig. 1. The pTTS system may take advantage of different pTTS templates to output different sentences of text in different voices, thereby providing output in the form of a multi-person conversation. Personal computer 110 generates the sound corresponding to the speech data, thereby enabling a

10 recipient interacting with personal computer 110 to hear the spoken message.

Referring to Fig. 3, an embodiment includes an author of a text message interacting with a first computer 120, and an intended recipient of the message interacting with a second computer 122. Computers 120 and 122 are coupled to data network 124 through Internet service provider 126 and Internet service provider 128, respectively. In alternative

15 embodiments, the data network may comprise the Internet, a company's internal data network, or a combination of several networks.

Server 130 couples to data network 124. Server 130 is a general purpose computer programmed to function as a web site. Server 130 also couples to storage device 132, such as a magnetic, optical, or magneto-optical storage device. Storage device 132 stores a

20 pTTS template 134 associated with the author, and may additionally store pTTS templates associated with other users. In an alternative embodiment, computer 120 transmits the author's pTTS template 134 to server 130 each time pTTS template 134 is needed, rather than storing pTTS template 134 on storage device 132.

515359_1

The author interacting with computer 120 generates text data intended for the recipient interacting with computer 122. Rather than transmitting the text data directly to computer 122, the text data is directed through data network 124 to server 130 for conversion to speech data. Conversion routine 136, executing in memory 138 of server 130, accepts the text

5   data and converts the text data to speech data with the author's pTTS template 134, using the process described in Fig. 1. The speech data thus contains information representing the voice characteristics of the author's speech template. Server 130 thereafter directs the speech data to computer 122. Server 130 may also send the original text data to computer 122, if desired. The recipient may listen to the speech message corresponding to the original text message with

10   software executing on computer 122, in the author's own voice or a voice selected by the author.

In an alternative embodiment, computer 120 sends the text file directly to computer 122 through data network 124. Computer 120 provides the necessary information for accessing the author's pTTS template 134 stored on storage 132 of server 130 to computer 122, thereby allowing the recipient to obtain speech data having characteristics of the author's voice.

15   The recipient interacting with computer 122 submits the text data to server 130 through data network 124, for conversion to speech data with conversion routine 136 and the author's pTTS template 134. Server 130 thereafter directs the speech data back to computer 122 for access by the recipient.

In another alternative embodiment, the text message is sent from computer 120 to

20   server 130. After converting the text data to speech data with conversion routine 136 and the author's pTTS template 134, server 130 returns the resulting speech data back to computer 120. Computer 120 sends the speech data directly to computer 122 through data network 124.

Referring to Fig. 4, in an alternative embodiment, storage device 140 coupled to

computer 120 stores the author's pTTS template 134. Alternatively, computer 120 downloads the author's pTTS template 134 from server 130 when necessary for conversion of text to speech. Conversion routine 136 executes in memory 142 of computer 120, for conversion of text data from the author into speech data. Therefore, computer 120 sends the speech data directly to

5   computer 122.

Referring to Fig. 5, in an alternative embodiment, storage device 144 coupled to computer 122 stores the author's pTTS template 134. Computer 120 separately sends the author's pTTS template 134 to computer 122. Alternatively, computer 122 downloads the author's pTTS template 134 from server 130. Conversion routine 136 executes in memory 146

10   of computer 122, for converting text data received from computer 120 into speech data. Therefore, computer 120 simply sends the text data to computer 122, which computer 122 converts to speech data if desired.

Referring to Fig. 6, in an alternative embodiment, server 130 is further coupled to public switched telephone network (PSTN) 148. Telephone 150 is also coupled to PSTN 148.

15   In one embodiment, PSTN 148 operates in a circuit switched manner, whereas data network 124 operates in a packet switched manner.

The embodiments illustrated herein describe computers coupled to a data network or coupled together through a data network. Coupling is defined herein as the ability to share information, either in real-time or asynchronously. Coupling includes any form of connection, either by wire or by means of electromagnetic or optical communications, and does not require that both computers are connected with the network at the same time. For example, a first and second computer are coupled together if a first computer accesses a network to send text data to an e-mail server, and the second computer retrieves such text data, or speech data associated

therewith, after the first computer has physically disconnected from the network.

The pTTS system described herein may provide a wide array of individualized services. For example, personalized templates are submitted with text to a known text-to-speech algorithm, thereby producing individualized speech from generic text. Therefore, a user of the

5   system may have a single pTTS template for use with text from a multitude of sources. Some of the uses of the pTTS system are discussed below.

Speech Announcements

In one embodiment, personal computer 110 of Fig. 2 is configured to operate as a

10 voice response system. For example, personal computer 110 is placed at a kiosk, and provides spoken delivery of stored information. As another example, personal computer 110 is coupled to the PSTN and configured to operate as a voice response system in response to user input provided via telephone key depression or speech. Voice response software is well-known. Examples of voice response systems are described by U.S. Patent No. 6,014,428, entitled "Voice

15 Templates For Interactive Voice Mail And Voice Response System", and U.S. Patent No. 5,125,024, entitled "Voice Response Unit", which are hereby incorporated by reference.

According to the present technique, the voice response software of personal computer 110 includes conversion routine 118, which is configured to use a pTTS template stored on storage 114. In one embodiment, the pTTS template represents the voice characteristics of the author. Alternatively, the pTTS template represents voice characteristics

20 characteristics of the author. Alternatively, the pTTS template represents voice characteristics selected by the author or the provider of the voice response system. For example, the system may select a pTTS template representing voice characteristics of a person similar to the user of the system, for example of the same gender or of a similar age. Alternatively, the system selects

11

a pTTS template predicted to elicit a certain response from the user, which may be based on marketing or psychological studies. Alternatively, the system allows the user to select which pTTS template to use.

5   The voice response system converts variable text messages to speech with a pTTS template. Some messages may contain both a variable portion and a fixed portion. One example of such message is "Your account balance is xx dollars and yy cents", where "xx" and "yy" are variable numerical values. In one embodiment, the entire text message comprising both the variable and fixed portions is submitted to the pTTS system for conversion to speech data. Alternatively, the fixed portions are prerecorded speech, and only the variable portions are

10   submitted as text to the pTTS system for conversion to speech data using the same voice that recorded the fixed portion of the message. A single audible message may be output by merging the prerecorded speech and generated speech data. In another embodiment, the entire text message is fixed text. Submitting such text to the pTTS system allows selecting the desired pTTS template based upon the factors as described above.

15

Film Dubbing

In another embodiment, personal computer 110 of Fig. 2 is configured to operate as part of a film editing system. Specifically, personal computer 110 operates to dub voices for films with foreign language subtitles. The pTTS templates of the actors are stored in storage

20   114, and used to produce speech data corresponding to the subtitles, thereby creating a multi-lingual soundtrack. In one embodiment, the lines of the actors are stored in a text file. An electronic code precedes each actor's lines, thereby identifying each portion of text with the correct actor. The code enables conversion routine 118 to select the correct pTTS template 116

515359_1

associated with the actor speaking a particular set of lines. The actors may need to produce

different templates for each language, due to the different pronunciation characteristics of words

in different languages. Timing information may be included in the text file to aid in the

production of speech data that is properly synchronized with the film. In an alternative

5   embodiment, a person's pTTS template may be used for different animated characters in

animated films.

Person-To-Person Spoken Messaging

In an alternative embodiment, computer 120 and computer 122 are each

10   configured with software for exchanging typed messages over data network 124, in a so-called

"instant message" format. Software that enables personal computers to exchange messages in

this manner is well known.

In the configuration shown in Fig. 3, the author types a text message using

computer 120 for delivery to computer 122. However, rather than sending the message directly

15   to computer 122, computer 120 directs the message through data network 124 to server 130.

Conversion routine 136 executing in memory 138 of server 130 converts the text data to speech

data, using the author's pTTS template 134, stored on storage 132. Server 130 thereafter directs

the speech data to computer 122. A person interacting with computer 122 may also act as the

initiator of a message, in which case such person's pTTS template is also stored on storage 132

20   of server 130. Messages directed to computer 120 are first directed to server 130 for conversion

to speech data using the appropriate pTTS template.

In the configuration shown in Fig. 4, the author types a text message using

computer 120 for delivery to computer 122. However, rather than sending the text message to a

13

centralized server, the message is converted to speech data by conversion routine 136 executing

in memory 142 of computer 120. The author's pTTS template 134 is stored on storage 140 of

computer 120, for access by conversion routine 136. Therefore, computer 120 sends the speech

data directly to computer 122 through data network 124. A person interacting with computer

5   122 may also act as the initiator of a message, in which case the message is converted to speech

data by the conversion routine executing in memory of computer 122, using the appropriate

pTTS template.

In the configuration shown in Fig. 5, the author types a text message using

computer 120, which is sent directly to computer 122 through data network 124. The author's

10   pTTS template 134 is stored on storage 144 of computer 122. Therefore, conversion routine 136

executing in memory 146 of computer 122 converts the text data to speech data. Alternatively,

computer 122 may direct the text data to server 130 for conversion to speech data using the

author's pTTS template 134 on storage 132 of server 130. Server 130 then redirects the speech

data back to computer 122. As in the other configurations, a person interacting with computer

15   122 may also act as the initiator of the message.

Chat Room Spoken Text

In an alternative embodiment, server 130 is operative to execute so-called Chat

software. In general, the Chat software enables a user to "enter" a chat room, view messages

20   input by other users who are in the chat room, and to type messages for display to all other users

in the chat room. The set of users in the chat room varies as users enter or leave.

Each Chat implementation architecture provides a Chat Client program and a

Chat Server program. The Chat Client program allows the user to input information and control

515359_1

which Chat Client users will receive such information. Chat Client user groupings, which may

be referred to as chat rooms or worlds, are the basis of the user control. A user controls which

Chat users will receive the typed information by becoming a member of the group that contains

the target users. A Chat user becomes a member of a group by executing a Chat Client "join

5   group" function. This function registers the Client's internet protocol (IP) address with the Chat

Server as a member of that group. Once registered, the Client can send and receive information

with all the other Clients in that group via the Chat Server. The exchange of information

between the Clients and Server is based on the "Internet Relay Chat" (IRC) protocol running

over separate input and output ports.

10   Fig. 7 illustrates a chat implementation architecture. Server 130 supports chat

group 152 and chat group 154. Other chat groups may be added. Users interacting through chat

client 156 and chat client 158 join chat group 152, and thereafter may communicate through chat

group 152 with the IRC protocol. Similarly, users interacting through chat client 160 and 162

join chat group 154, and thereafter may communicate through chat group 154 with the IRC

15   protocol.

According to the present technique, at least one user in the chat room has access

to a computer operative to generate speech with the user's pTTS template.

In the configuration shown in Fig. 3, server 130 acts as the chat room. Storage

132 stores the pTTS templates for each user in the chat room. A user's pTTS template is

20   transferred to server 130 when the user signs in to the chat room. Server 130 stores the pTTS

templates of frequent users, to avoid the necessity of submitting the pTTS template each time a

user signs in. Thereafter, as each user submits text data to the chat room, conversion routine 136

executing in memory 138 of server 130 converts the text data to speech data using the

15

submitter's pTTS template. Therefore, each user can access messages from other users having the voice characteristics of the corresponding user. The server may also provide text messages, in the event that some users do not provide a pTTS template. The personalized speech may be delivered as an audio file in ".wav" format or other suitable format. Alternatively, the

5   personalized speech may be delivered from server 130 as streaming audio.

In the configuration shown in Fig. 4, server 130 acts as the chat room. However, the pTTS template 134 of each user is stored on storage 140 of the user's computer 120. In an alternative embodiment, the user's pTTS template 134 is downloaded from server 130 as the user enters the chat room. As the user leaves the chat room, server 130 notifies the user's computer

10   120 that the pTTS template is no longer needed, so that it may be deleted from storage 140. Each user, therefore, sends speech data directly to the chat room, as opposed to text data.

In the configuration shown in Fig. 5, server 130 acts as the chat room. Server 130 stores the pTTS template of each user in storage 132. When a user enters the chat room, the user downloads the pTTS templates of each user in the chat room, and stores the pTTS templates on

15   storage 144 of the user's computer 122. Messages are submitted to server 130 in text format, and read by the user's computer 122 in text format. However, when computer 122 receives messages typed by another user in the chat room, such as a user interacting with computer 120, computer 122 generates speech corresponding to the text of the message using the author's pTTS template 134 stored on storage 144.

20   In an alternative embodiment, personalized speech is delivered to a telephone-only participant in the chat room, interacting through telephone 164. Automated speech recognition (ASR) functions 166 and pTTS functions interface with the standard Chat architecture via Chat Proxy 168. Chat Proxy 168 establishes the Chat session with the Chat

Server, joins the appropriate group, and establishes an input session with ASR 166 and an output

session with the pTTS functions. ASR 166 converts the phone speech to text and sends the

output to Chat Proxy 168. Chat Proxy 168 takes the text stream from ASR 166 and delivers it to

the Chat Server input port using IRC. Chat Proxy 168 also converts the IRC stream from the

5   Chat Server output port into the original typed text and delivers it to the pTTS function where the

text is played to the phone user in the Chat Client user's voice.

Spoken Electronic Mail

Electronic mail systems having a text-to-speech front-end that allows a user to

10   retrieve their electronic mail using a telephone are known. However, in an embodiment of the

present invention, a user may listen to electronic mail in the author's own voice. For example, a

parent that is away from home may send an e-mail message to a child, who is then able to listen

to the message in the parent's own voice.

Referring to Fig. 6, let it be assumed that the user of computer 120 composes an

15   electronic mail message, indicates a preferred delivery time, and also indicates that it is to be

delivered via speech to a particular telephone number, such as the telephone number associated

with telephone 150. The user of computer 120 sends this message via ISP 126 and data network

124 to server 130. Server 130 stores the message in storage 132. At the preferred delivery time,

server 130 retrieves the message from storage 132, and also retrieves the author's pTTS template

20   134 from storage 132. It will be appreciated by those skilled in the art that the message and the

pTTS template may be stored on different storage devices. Server 130 uses the author's

retrieved pTTS template 134 to generate speech corresponding to the retrieved message.

Specifically, conversion routine 136 executing in memory 138 of server 130 converts the text

17

message to speech data. Server 130 then places a telephone call using PSTN 148 to telephone 150 and delivers the personalized speech.

In an alternative embodiment, spoken electronic mail is implemented as person-to-person spoken messaging, as described above with reference to Figs. 3-5.

5

Shared Space Objects

A "shared space" is a location on the Internet where members of a group can store objects, so that other members of the group can access those objects. A chat room is an example of a real-time shared space location, although a shared space provides additional flexibility by allowing storage of objects for future access. Such Internet hosting systems that allow users to upload objects and control object access are known.

In an embodiment of the present invention, a user creates an object and associates the user's pTTS template with the object. The object-pTTS template association may be to the object (text file), and/or an object description (text file describing the object). The user uploads the object and the user's associated pTTS template to the Internet site shared space. Thereafter, when another user with permission to access the shared object accesses that object, a pTTS enabler provides the user the option to hear the speech associated with the text. The pTTS enabler may be invoked automatically, or on demand. If the user selects to hear the message, a conversion routine converts the text data to speech data using the corresponding pTTS template.

20   In one embodiment, a shared space object comprises biographical information describing a user, in text format. Therefore, by converting the text data to speech data with the user's pTTS template, other users may hear the biographical description in the user's own voice.

In other embodiments, shared space objects may include classified ads, resumes, personal web

sites, or other personal information.

Spoken Telephone Call Notice

5   U.S. Patent No. 5,805,587, the disclosure of which is hereby incorporated by

reference, describes a facility to alert a subscriber whose telephone is connected to the Internet of

a waiting call, the alert being delivered via the Internet. A waiting call is forwarded from the

PSTN to a services platform that sends the alert to the subscriber via the Internet. If requested by

the subscriber, the platform may then forward the telephone call to the subscriber via the Internet

10   without interrupting the subscriber's Internet connection.

         Referring to Fig. 6, the user of telephone 150 is assumed to be calling the user of

computer 120. The user of computer 120 is assumed to have a telephone (not shown) that is not

coupled to PSTN 148, because the user of computer 120 is instead using the telephone line to

connect to ISP 126. Server 130 operates as the services platform described in U.S. Patent No.

15   5,805,587, and delivers a message via data network 124 and ISP 126 to computer 120 that a call

from telephone 150 is waiting. The user of computer 120 composes a textual message, or

retrieves an already composed textual message, for delivery to the user of telephone 150, and

sends the message from computer 120 via ISP 126 and data network 124 to server 130. Server

130 retrieves the pTTS template 134 for the user of computer 120 from storage 132, generates

20   speech corresponding to the message using conversion routine 136 executing in memory 138,

and delivers the personalized speech via PSTN 148 to telephone 150.

Personalized Speech For Software Applications

In another embodiment, personal computer 110 of Fig. 2 is configured to operate as a pTTS system in cooperation with a software application. The software application submits text data to conversion routine 118 executing in memory 112, for conversion to speech data. The speech data is output to a user as audio information through speakers coupled to personal computer 110. Conversion routine 118 operates as an independent program, which may be accessed by various software applications for conversion of text data to speech data. Alternatively, conversion routine 118 is integrated with the software application requiring text-to-speech services.

In one embodiment, the software application comprises a learning program that provides an interactive teaching session with a user. Learning programs providing pre-recorded audio output are known. However, the pTTS system provides personalized audio output in place of such pre-recorded audio. Specifically, the learning program submits text data to conversion routine 118, which converts the text data to speech data having characteristics of a specified voice. The voice could be of a parent or teacher, thereby personalizing the learning experience.

In another embodiment, the text of a book or article is submitted to conversion routine 118 for conversion to speech data. A parent may include his or her speech template in storage 114, permitting a child to hear the book or article read in the parent's own voice.
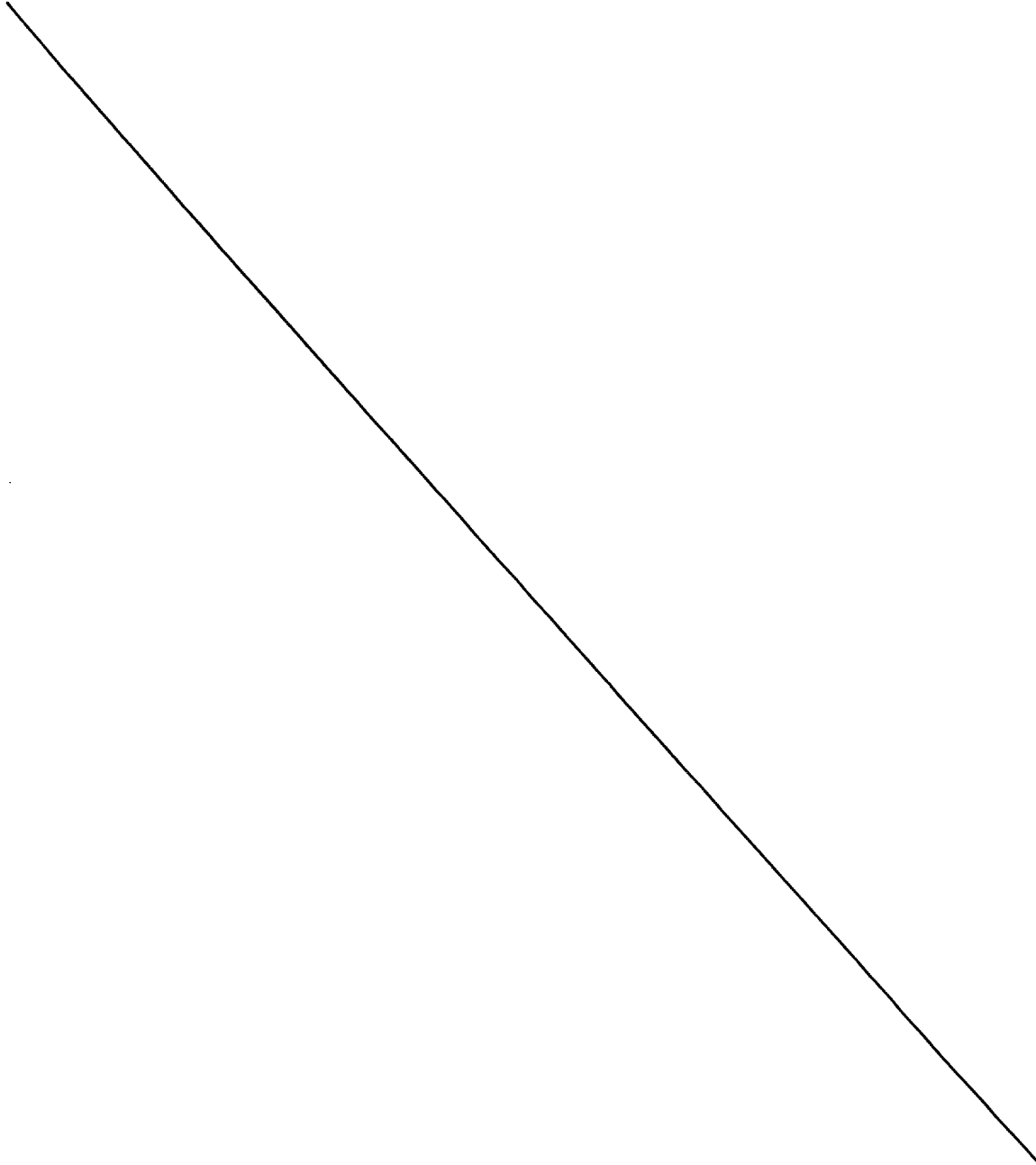
In another embodiment, the pTTS system is implemented in a device such as a children's toy, which is capable of executing conversion routine 118 and storing pTTS template 116. A pTTS template is loaded into the device, thereby providing personalized speech output.

Although illustrative embodiments of the present invention and various modifications thereof have been described in detail herein with reference to the accompanying

20

drawings, it is to be understood that the invention is not limited to these precise embodiment and

the described modifications, and that various changes and further modifications may be effected

therein by one skilled in the art without departing from the scope or spirit of the invention as

defined in the appended claims.

21

CLAIMS

1.   A method for converting text to speech comprising:

receiving data representing a textual message, said message being directed from an author to a recipient;

receiving information identifying an individual;

retrieving a speech template comprising information representing characteristics of said individual's voice; and

converting said data representing said textual message to speech data, said speech data representing a spoken form of said textual message having the characteristics of said individual's voice.

2.   The method according to claim 1 wherein said author interacts with a first computer and said recipient interacts with a second computer coupled to said first computer through a data network.

3.   The method according to claim 2 wherein said speech template is provided at a central location coupled to said first computer and said second computer.

4.   The method according to claim 3 further comprising receiving said data representing the textual message at said central location from said first computer.

5.   The method according to claim 4 further comprising transmitting said speech data to said second computer from said central location.

6.   The method according to claim 4 further comprising transmitting said speech data to said first computer from said central location.

7.   The method according to claim 6 further comprising transmitting said speech data to said second computer from said first computer.

8.   The method according to claim 3 further comprising receiving said data representing the textual message at said central location from said second computer.

9.   The method according to claim 8 further comprising transmitting said speech data to said second computer from said central location.

10.   The method according to claim 2 wherein said speech template is provided at said first computer.

11.   The method according to claim 10 further comprising transmitting said speech data or said speech template to said second computer from said first computer.

12.   The method according to claim 2 wherein said speech template is provided at said second computer.

13.   The method according to claim 12 further comprising receiving said data

representing the textual message at said second computer from said first computer.

14.   The method according to claim 2 wherein said first computer and said second computer are configured to communicate in an instant messaging format.

15.   The method according to claim 2 wherein said first computer and said second computer are coupled to a server configured to operate chat room software, said data representing the textual message comprising text input to said chat room.

16.   The method according to claim 15 wherein said server stores speech templates for users of said chat room.

17.   The method according to claim 2 wherein said first computer and said second computer are coupled to a server adapted to store and provide access to a shared space object, said shared space object being associated with said textual message.

18.   The method according to claim 1 wherein said author comprises computer executable program code designed to generate text in response to input from said recipient.

19.   The method according to claim 18 wherein a computer executing said program code couples to a public switched telephone network.

20.   The method according to claim 19 wherein said input from said recipient

515359_1

comprises telephone key depression or speech.

21.   The method according to claim 18 further comprising selecting said individual based on attributes of said recipient.

22.   The method according to claim 21 wherein said attributes comprise age or gender.

23.   The method according to claim 1 where said recipient interacts with a telephone coupled to a telephone network and said author interacts with a computer coupled to said telephone network through a data network.

24.   The method according to claim 23 further comprising directing said speech data to said telephone network through said data network.

25.   The method according to claim 23 further comprising transmitting a notification to said author when said recipient is unable to connect with a telephone of said author.

26.   The method according to claim 25 further comprising receiving said data representing the textual message in response to said notification message.

27.   The method according to claim 1 wherein said data representing the

515359_1

textual message comprises a variable portion of a message having both a variable portion and a

fixed portion.

28.   The method according to claim 27 wherein said data representing the

textual message further comprises said fixed portion.

29.   The method according to claim 27 wherein said fixed portion is

prerecorded speech of said individual.

30.   The method according to claim 1 wherein said data representing the

textual message comprises an e-mail message.

31.   A text to speech conversion system comprising:

a memory that stores executable program code;

a processor that executes said program code;

a storage device that stores a speech template comprising information

representing characteristics of said individual's voice, said individual being identified by

identification data; and

wherein said program code is executable to convert text data to speech data, said

text data representing a textual message that is directed from an author to a recipient, and said

speech data representing a spoken form of said text data having the characteristics of said

individual's voice.

32.   The system according to claim 31 wherein said author interacts with a first computer and said recipient interacts with a second computer coupled to said first computer through a data network.

33.   The system according to claim 32 further comprising a centralized computer that includes said processor wherein said centralized computer couples to said first computer and said second computer.

34.   The system according to claim 33 wherein said centralized computer further comprises a communications port adapted to receive said text data from said first computer or said second computer.

35.   The system according to claim 34 wherein said centralized computer is adapted to transmit said speech data to said first computer or said second computer.

36.   The system according to claim 33 wherein said centralized computer is configured to operate chat room software.

37.   The system according to claim 36 wherein said centralized computer stores speech templates for users of said chat room, said text data comprising text input to said chat room.

38.   The system according to claim 33 wherein said centralized computer is

515359_1

adapted to store and provide access to a shared space object, said shared space object being associated with said text data.

39.   The system according to claim 32 wherein said first computer comprises said storage device.

40.   The system according to claim 39 wherein said first computer further comprises a communications port that transmits said speech data or said speech template to said second computer.

41.   The system according to claim 32 wherein said second computer comprises said storage device.

42.   The system according to claim 41 wherein said second computer further comprises a communications port that receives said text data.

43.   The system according to claim 32 wherein said first computer and said second computer are configured to communicate in an instant messaging format.

44.   The system according to claim 31 wherein said author comprises voice-response program code designed to generate text in response to input from said recipient.

45.   The system according to claim 44 wherein a computer comprising said

515359_1

processor couples to a public switched telephone network.

46.   The system according to claim 45 wherein said input from said recipient

comprises telephone key depression or speech.

47.   The system according to claim 44 further comprising selection program

code that selects said individual based on attributes of said recipient.

48.   The system according to claim 47 wherein said attributes comprise age or

gender.

49.   The system according to claim 31 where said recipient interacts with a

telephone coupled to a telephone network and said author interacts with a computer coupled to

said telephone network through a data network.

50.   The system according to claim 49 further comprising directing said speech

data to said telephone network through said data network.

51.   The system according to claim 49 further comprising notification program

code designed to transmit a notification to said author when said recipient is unable to connect

with a telephone of said author.

52.   The system according to claim 51 wherein said text data comprises a

textual message in response to said notification message.

53.   The system according to claim 31 wherein said text data is a variable portion of a message having both a variable portion and a fixed portion.

54.   The system according to claim 53 wherein said text data further comprises said fixed portion.

55.   The system according to claim 53 wherein said fixed portion is prerecorded speech of said individual.

56.   The system according to claim 31 wherein said text data comprises an e-mail message.

57.   A text to speech conversion system comprising:

means for receiving data representing a textual message, said message being directed from an author to a recipient;

means for receiving information identifying an individual;

means for retrieving a speech template comprising information representing characteristics of said individual's voice; and

means for converting said data representing said textual message to speech data, said speech data representing a spoken form of said textual message having the characteristics of said individual's voice.

515359_1

58.   The system according to claim 57 wherein said author comprises computer executable program code designed to generate text in response to input from said recipient.

59.   The system according to claim 58 further comprising means for receiving telephone key depression or speech from said recipient as input.

60.   The system according to claim 58 further comprising means for selecting said individual based on attributes of said recipient.

61.   The system according to claim 57 further comprising means for delivering said speech data to a telephone network wherein said recipient interacts with a telephone coupled to said telephone network.

62.   The system according to claim 61 further comprising means for transmitting a notification to said author when said recipient is unable to connect with a telephone of said author.

63.   The system according to claim 62 further comprising means for receiving said data representing the textual message in response to said notification message.

64.   The system according to claim 57 wherein said data representing the textual message comprises a variable portion of a message having both a variable portion and a

515359_1

fixed portion.

65.   The system according to claim 64 wherein said data representing the textual message further comprises said fixed portion.

66.   The system according to claim 64 wherein said fixed portion is prerecorded speech of said individual.

67.   The system according to claim 57 wherein said data representing the textual message comprises an e-mail message.

68.   An article of manufacture comprising:

a computer readable medium having computer usable program code embodied therein, said computer usable program code containing executable instructions that when executed, cause a computer to perform the steps of:

receiving data representing a textual message, said message being directed from an author to a recipient;

receiving information identifying an individual;

retrieving a speech template comprising information representing characteristics of said individual's voice; and

converting said data representing said textual message to speech data, said speech data representing a spoken form of said textual message having the characteristics of said individual's voice.

515359_1

69.   The article according to claim 68 wherein said author interacts with a first computer and said recipient interacts with a second computer coupled to said first computer through a data network.

70.   The article according to claim 69 wherein said speech template is provided at a central location coupled to said first computer and said second computer.

71.   The article according to claim 69 wherein said speech template is provided at said first computer.

72.   The article according to claim 69 wherein said speech template is provided at said second computer.

73.   The article according to claim 69 wherein said first computer and said second computer are configured to communicate in an instant messaging format.

74.   The article according to claim 69 wherein said first computer and said second computer are coupled to a server configured to operate chat room software, said data representing the textual message comprising text input to said chat room.

75.   The article according to claim 74 wherein said server stores speech templates for users of said chat room.

76.   The article according to claim 69 wherein said first computer and said second computer are coupled to a server adapted to store and provide access to a shared space object, said shared space object being associated with said textual message.

77.   The article according to claim 68 wherein said author comprises computer executable program code designed to generate text in response to input from said recipient.

78.   The article according to claim 77 wherein a computer executing said program code couples to a public switched telephone network.

79.   The article according to claim 78 wherein said input from said recipient comprises telephone key depression or speech.

80.   The article according to claim 77 wherein said information identifying an individual includes age or gender attributes.

81.   The article according to claim 68 where said recipient interacts with a telephone coupled to a telephone network and said author interacts with a computer coupled to said telephone network through a data network.

82.   The article according to claim 68 wherein said data representing the textual message comprises a variable portion of a message having both a variable portion and a

515359_1

fixed portion.

83. The article according to claim 82 wherein said data representing the textual message further comprises said fixed portion.

84. The article according to claim 82 wherein said fixed portion is prerecorded speech of said individual.

85. The article according to claim 68 wherein said data representing the textual message comprises an e-mail message.

86. A method for generating speech data for a voice response system comprising:

receiving input from a recipient;

generating a text message that provides a response to said input;

selecting a speech template comprising information representing characteristics of a voice based at least in part on attributes of said recipient; and

converting said text message to speech data, said speech data representing a spoken form of said textual message having the characteristics of said voice.

87. The method according to claim 86 wherein said input from said recipient comprises telephone key depression and speech.

515359_1

88.   The method according to claim 86 further comprising delivering said speech data to said recipient through a public switched telephone network.

89.   The method according to claim 86 where said attributes comprise age or gender.

90.   A method for converting chat room text to speech comprising:

storing a plurality of speech templates, each speech template comprising information representing characteristics of a chat room participant's voice;

receiving said chat room text from an author, said author being a chat room participant;

retrieving a speech template comprising information representing characteristics of said author's voice from said plurality of speech templates; and

converting said chat room text to speech data, said speech data representing a spoken form of said textual message having the characteristics of said author's voice.

91.   A method for providing spoken electronic mail comprising:

receiving an electronic text message from an author of said message, said message addressed to a recipient;

retrieving a speech template comprising information representing characteristics of said author's voice;

converting said text message to speech data, said speech data representing a spoken form of said textual message having the characteristics of said author's voice; and

directing said speech data to said recipient.

92.   The method according to claim 91 further comprising directing said

speech data to said recipient over a public switched telephone network.

93.   A method for providing speech output from a software application,

comprising:

receiving text data from said software application;

receiving information identifying an individual;

retrieving a speech template comprising information representing characteristics
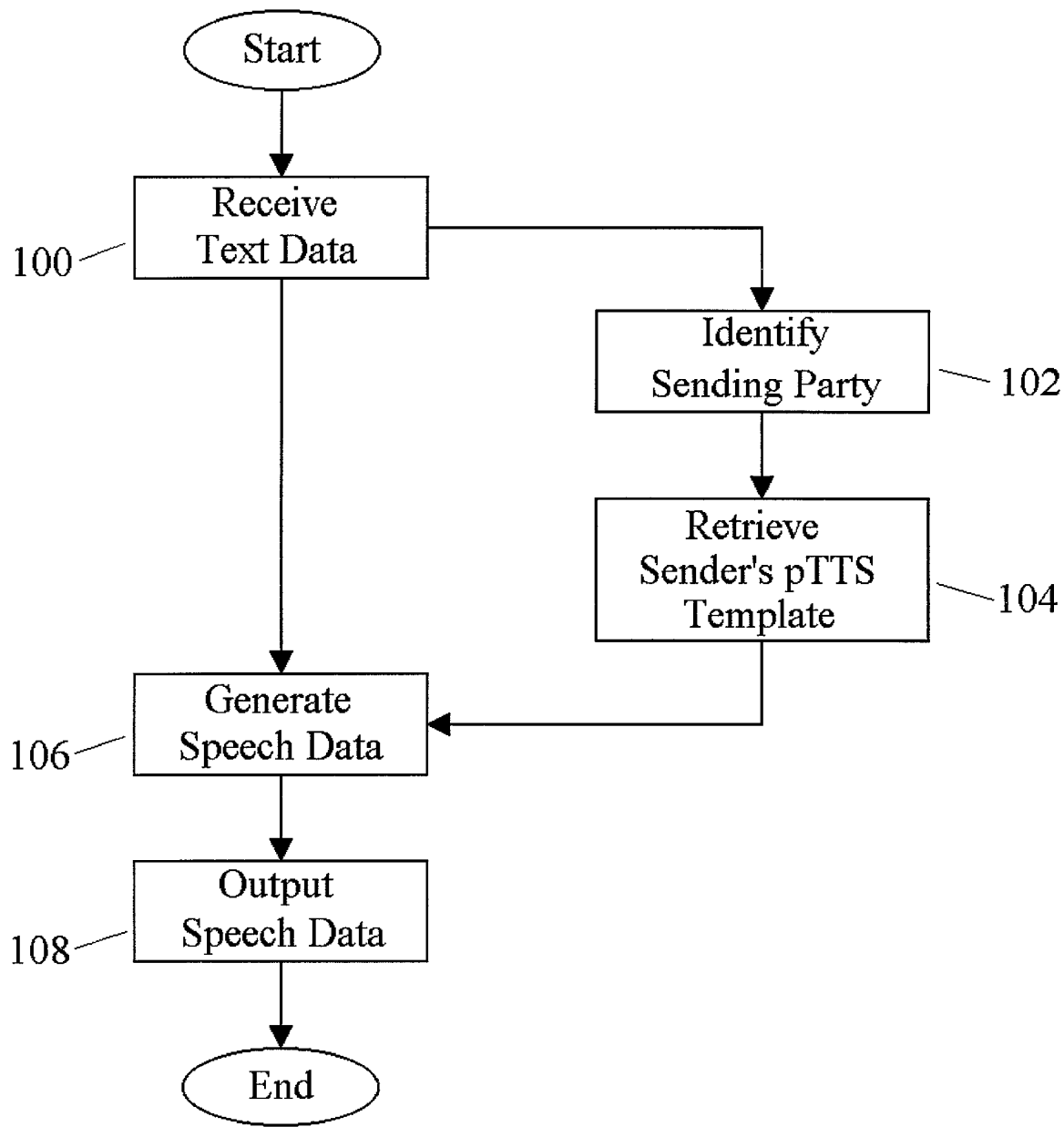
of said individual's voice;

converting said text data to speech data, said speech data representing a spoken

form of said text data having the characteristics of said individual's voice; and

supplying said speech data to an output device for output to a user as audio

information.

94.   The method according to claim 93 wherein the software application

comprises an interactive learning program.

515359_1

ABSTRACT
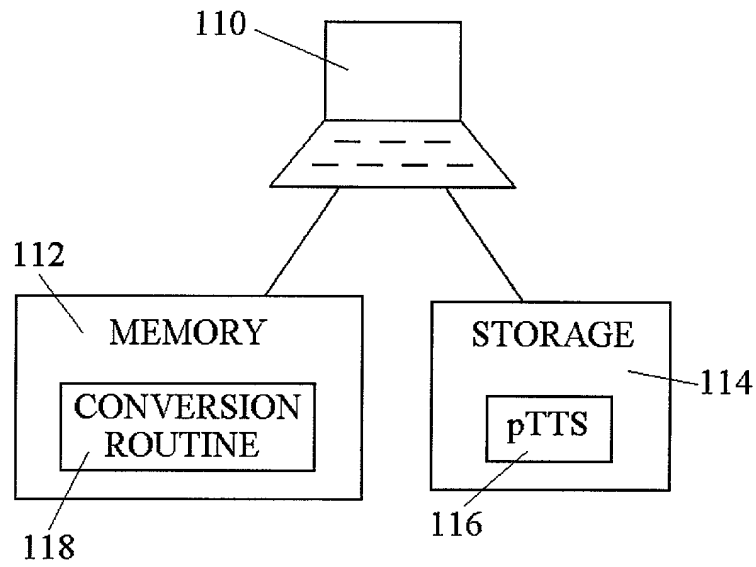
A personalized text-to-speech (pTTS) template is used to provide various services including speech announcements, film dubbing, Internet person-to-person spoken messaging, Internet chat room spoken text, spoken notice of an incoming telephone call to a subscriber using the Internet, spoken electronic mail and Internet shared spaces having objects intended for spoken presentation.
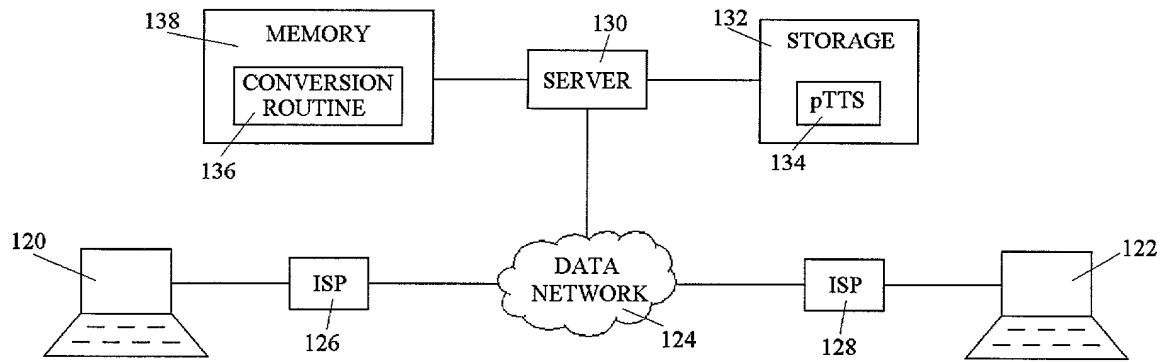
515359_1

```
        ┌─────────┐
        │  Start  │
        └────┬────┘
             │
             ▼
   ┌──────────────────┐
100│     Receive      ├──────────────┐
   │    Text Data     │              │
   └────────┬─────────┘              ▼
            │              ┌──────────────────┐
            │              │     Identify     │
            │              │  Sending Party   │──── 102
            │              └────────┬─────────┘
            │                       │
            │                       ▼
            │              ┌──────────────────┐
            │              │     Retrieve     │
            │              │  Sender's pTTS   │──── 104
            │              │     Template     │
            │              └────────┬─────────┘
            ▼                       │
   ┌──────────────────┐            │
106│     Generate     │◄───────────┘
   │   Speech Data    │
   └────────┬─────────┘
            │
            ▼
   ┌──────────────────┐
108│     Output       │
   │   Speech Data    │
   └────────┬─────────┘
            │
            ▼
        ┌─────────┐
        │   End   │
        └─────────┘
```
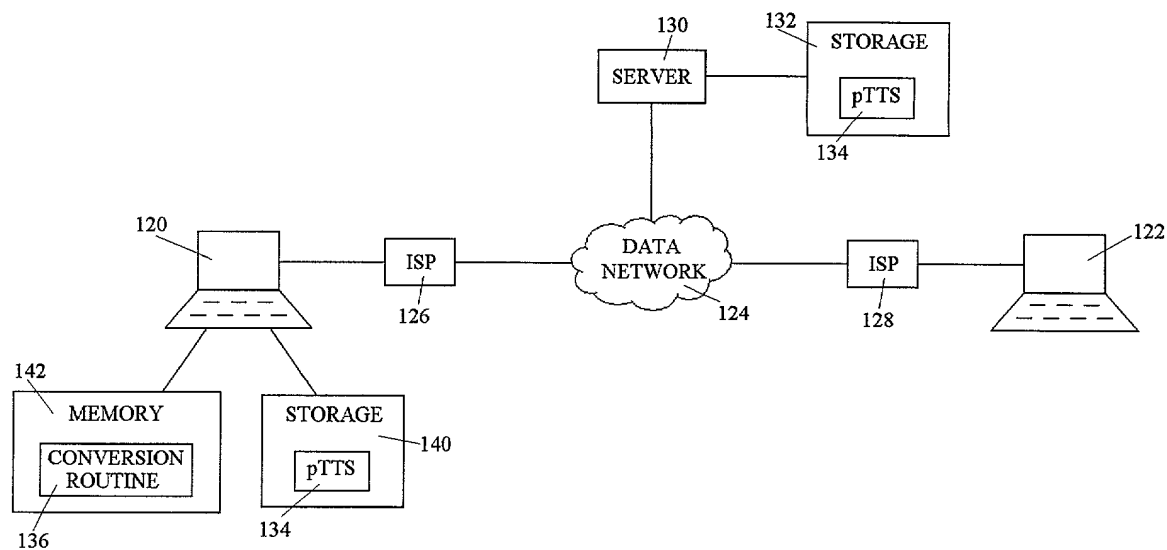
FIG. 1

515359_1

110

112

MEMORY

CONVERSION
ROUTINE

118

STORAGE

pTTS

114

116

FIG. 2

138 MEMORY

CONVERSION
ROUTINE

136

130

SERVER

132 STORAGE

pTTS

134

120

ISP

126

DATA
NETWORK

124

ISP

128

122

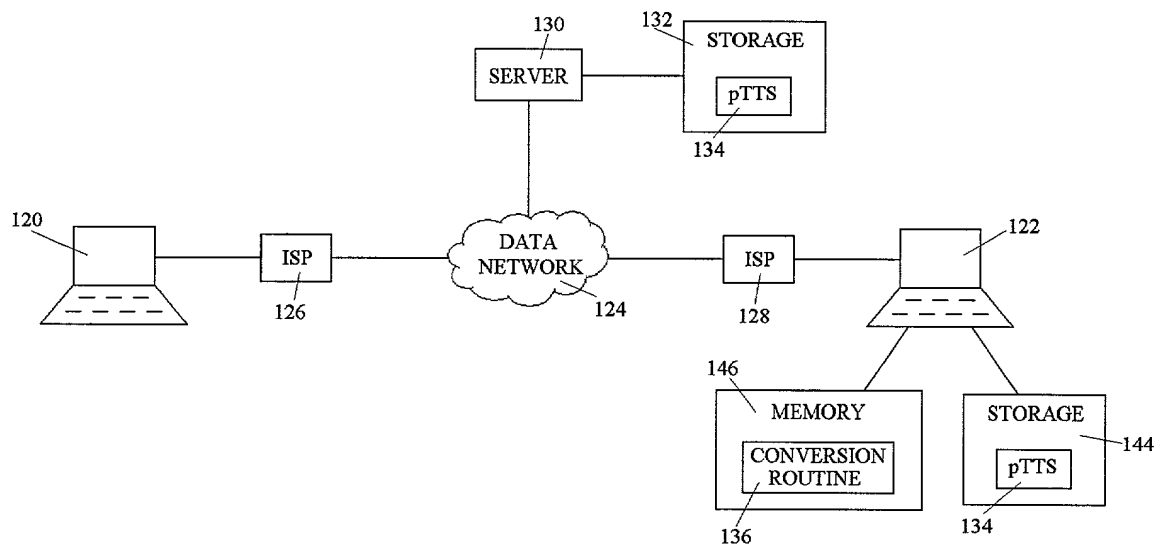**FIG. 3**
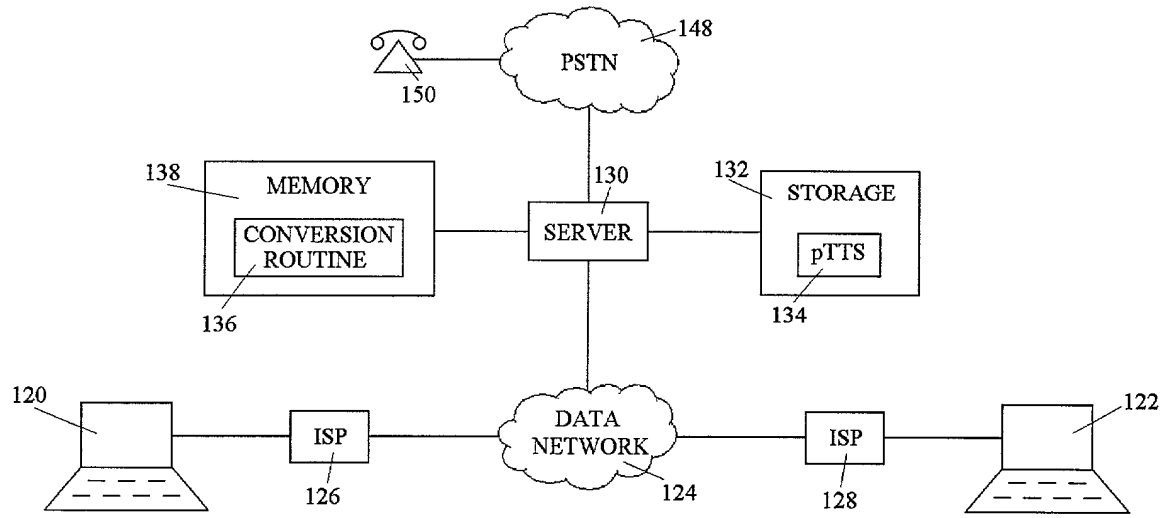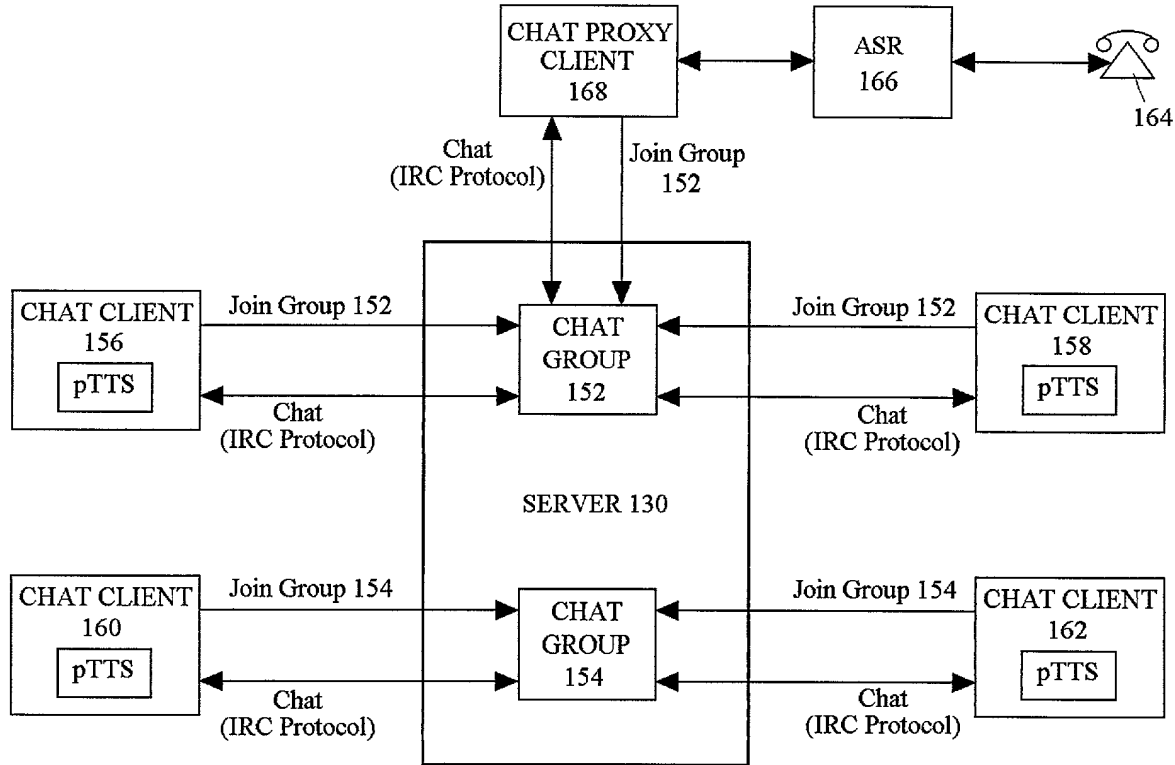
515359_1

FIG. 4

FIG. 5

515359_1

FIG. 6

FIG. 7

IN THE UNITED STATES
PATENT AND TRADEMARK OFFICE

Declaration and Power of Attorney

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled **Personalized Text-To-Speech Services,** the specification of which is attached hereto.

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by an amendment, if any, specifically referred to in this oath or declaration.

I acknowledge the duty to disclose all information known to me which is material to patentability as defined in Title 37, Code of Federal Regulations, 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, 119(a-d) or 365(a-b) of PCT or foreign application(s) for patent or inventors' certificate listed below or priority benefits under 119(e) of any United States provisional application(s) listed below and have also identified below any foreign application for patent or inventors' certificate having a filing date before that of the application on which priority is claimed:

None

I hereby claim the benefit under Title 35, United States Code, 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, 112, we acknowledge the duty to disclose all information known to us to be material to patentability as defined in Title 37, Code of Federal Regulations, 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

None

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

I hereby appoint the following attorney(s) with full power of substitution and revocation, to prosecute said application, to make alterations and amendments therein, to receive the patent, and to transact all business in the Patent and Trademark Office connected therewith:

Samuel H. Dworetsky	(Reg. No. 27873)
Thomas A. Restaino	(Reg. No. 33444)
Michele L. Conover	(Reg. No. 34962)
Cedric G. DeLaCruz	(Reg. No. 36498)
Rohini K. Garg	(Reg. No. 45272)
Benjamin S. Lee	(Reg. No. 42787)
Robert B. Levy	(Reg. No. 28234)
Susan E. McHale	(Reg. No. 35948)
Jeffrey M. Navon	(Reg. No. 32711)
Alfred G. Steinmetz	(Reg. No. 22971)

I also appoint Christopher A. Hughes (Reg. No. 26914, David V. Rossi (Reg. No. 36659), and David M. La Bruno (Reg. No. 46266) of Morgan & Finnegan as associate attorneys, with full power to prosecute said application, to make alterations and amendments therein, and to transact all business in the Patent and Trademark Office connected therewith.

Please address all correspondence to Mr. S. H. Dworetsky, AT&T Corp., P.O. Box 4110, Middletown, New Jersey 07748. Telephone calls should be made Michele L. Conover at 908-221-5773.

Full name of 1st joint inventor: Edmund Gale Acker

Inventor's signature _Edmund Gale Acker_   Date _6/29/2000_

Residence: Colts Neck, Monmouth County, New Jersey

Citizenship: United States of America

Post Office Address: 132 Hockhockson Road
                              Colts Neck, New Jersey   07722

Full name of 2nd joint inventor: Frederick Murray Burg

Inventor's signature _~Frederick Murray Burg~_____ Date _6/1/00_

Residence: West Long Branch, Monmouth County, New Jersey

Citizenship: United States of America

Post Office Address: 62 Golf Street
                                  West Long Branch, New Jersey   07764